CHAPTER 8
QUESTIONING COSMOLOGY – OUTSTANDING PROBLEMS ABOUT THE UNIVERSE

8.1 Introduction

> 'Don't let me catch anyone talking about the Universe in my department.'
>
> Lord Rutherford

Rutherford – the discoverer of the atomic nucleus – was a practical man. Perhaps one of the greatest experimental physicists of all time, he was profoundly sceptical of notions that were not grounded in hard experimental evidence. We can guess that he might not have been happy with some of the ideas of modern cosmology! But, however challenging the concepts of cosmology – and ideas about space being curved or expanding, and the vacuum possessing energy are certainly challenging – there can be no doubt that Rutherford would have been deeply impressed and intrigued by the vast amount of observational evidence that cosmologists have now acquired. His main interest, however, might well have been in those areas of cosmology where there are still clear gaps in our knowledge; the areas where work remains to be done and where new insights can be expected to arise.

The last few chapters have been largely concerned with the development and testing of models of the Universe. By a model we mean some simplified representation of the real world that helps us to understand reality by focusing on some specific aspects of it. The model should be simple, but not so simple that the phenomena of interest are inadequately represented. The Earth's orbit around the Sun, for example, can be modelled by a circle. This is an adequate model for explaining the occurrence of certain annual events, but a better model, such as an ellipse, is required to explain finer details such as the precise timing of those events. Neither model fully represents reality, but both are useful within their own ranges of validity, and the greater precision of the elliptical model is bought at the price of greater mathematical complexity.

- ■ In what important way did the FRW cosmological models of Chapter 5 simplify reality?
- ❑ They treated the contents of the Universe as a simple uniform fluid, the properties of which could be specified at any time by a density $\rho(t)$ and a pressure $p(t)$. Small-scale departures from uniformity, such as stars and galaxies, or even whole clusters of galaxies, were ignored.

- ■ In what important way did the big bang model of Chapter 6 improve on this, and how was the modelling further extended in Chapter 7?
- ❑ In Chapter 6 the matter in the Universe was treated more realistically by taking account of the variety of interacting particles that it contains, and by acknowledging the presence of density fluctuations. In Chapter 7 a spectrum of density fluctuations was considered (characterized by the parameters A and n_s) and related to the observed anisotropies in the cosmic microwave background radiation.

Despite its conceptual difficulty, the big bang is now widely accepted as the best available model of how the Universe evolved into its present state. However, it is quite clear that the model, as it has been presented in this book, and insofar as it is accepted by the majority of cosmologists, is still inadequate in several important respects. There are a number of major questions about our Universe that the standard big bang model does not address. This does not mean the model is wrong, but it does indicate deficiencies in certain areas and the model may need to be extended in those areas. This chapter concerns some of these outstanding problems, and considers the ways in which the standard big bang model might be extended to provide a more adequate account of reality.

The questions we shall consider are these:

Problem 1: What is the dark matter? (Section 8.2)

Problem 2: What is the dark energy? (Section 8.3)

Problem 3: Why is the Universe so uniform? (Section 8.4.1)

Problem 4: Why does the Universe have a flat ($k = 0$) geometry? (Section 8.4.2)

Problem 5: Where did the structure come from? (Section 8.5)

Problem 6: Why is there more matter than antimatter? (Section 8.6)

Problem 7: What happened at $t = 0$? (Section 8.7)

Problem 8: Why is the Universe the way it is? (Section 8.8)

8.2 The nature of dark matter

As you have seen many times in this book, the visible matter in the Universe – the stuff of stars and nebulae – accounts for only a small fraction of the whole. By recent estimates about 85% of the matter in the Universe is dark matter, and its nature is still a mystery. This is hardly satisfactory! So far you have been asked to accept the existence of dark matter without really enquiring what it is, but the time has now come to address that question – Problem 1 in our list – head on.

To prepare for that, we will review some of what you have already learned about dark matter.

- What do you understand by baryonic and non-baryonic dark matter?
- Baryonic dark matter is non-luminous matter in which most of the mass is attributable to baryons, most probably neutrons and protons. Non-baryonic dark matter is non-luminous matter made from something else.

In Chapter 6 you saw that the physics of the early Universe, especially the nucleosynthesis of the elements, puts constraints on the density of baryonic matter. No more than about 15% of the matter in the Universe can be baryonic. If we can trust the physics, then some of the dark matter may be baryonic but most of it has to be non-baryonic. Figure 8.1 summarizes our current understanding of the composition of the Universe.

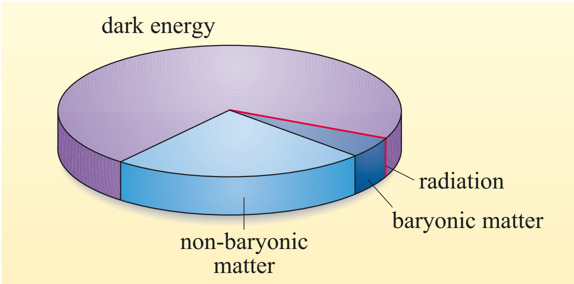


Figure 8.1 The contributions to the total (energy) density of the Universe from various sources. According to recent estimates, about 73% of the density of the Universe is currently due to dark energy, and about 27% is due to matter of all types. Only 4.4% of the total density is due to baryonic matter (i.e. roughly 15% of the matter). The contribution from radiation is only about 0.005%.

What can the dark matter be? Two broad classes of candidates have been proposed and they have been given the colourful names of MACHOs and WIMPs.

8.2.1 MACHOs

An obvious possibility, first mentioned in Chapter 1, is that at least some of the dark matter is simply normal matter – baryonic matter – that we cannot see and cannot detect other than by its gravitational influence. Stars are visible because they glow; if they did not we would regard them as dark matter.

QUESTION 8.1

Make a list of all the *known* types of astronomical object that could make up the dark matter in the halo of our own Galaxy.

In answering Question 8.1 you may have noticed that these dark objects fall into one of two categories – they are either *stellar remnants* or they are bodies that have masses lower than that of main sequence stars. All such objects are called **MACHOs**, a name that stands for ‘massive astrophysical compact halo object’. (Astrophysicists are known for their occasional lapses into whimsy.) MACHOs, if they exist, are simply ‘familiar’ objects that happen to emit little or no radiation of their own and have therefore evaded detection except by their gravitational effect.

Some MACHOs (dead stars, etc.) might reasonably be expected to exist, so an important question to answer is whether the dark matter in our Galaxy – and by extension that in other galaxies and clusters of galaxies – can be *entirely* accounted for by MACHOs or whether something else is needed. One approach is a theoretical one. From what we know about stellar evolution and the age of the Galaxy we can estimate how many of these different kinds of objects might have been formed and work out their total contribution to the cosmic density. Helpful though it is, that approach is not as compelling as an observational one. Is there any way we can actually detect and count MACHOs?

Yes, there is. In Chapter 4 you saw how gravitational lensing could be used to study the distribution of dark matter in a cluster of galaxies. A similar technique can be used to search for MACHOs in the Milky Way. Dark objects in the Galaxy will occasionally pass in front of background stars, bending their light and causing them to brighten. Even a planet-size object can act as a gravitational lens. This smaller scale phenomenon is known as **gravitational microlensing**. All we need to do is look at a sufficiently distant star and wait for it to brighten as a MACHO passes in front of it.

■ What problems do you see with this idea?

❑ Since both MACHOs and stars will have very small angular sizes, lensing events are likely to be rare. You would have to monitor a very large number of stars to stand a chance of seeing a single event. Also, you would have to be sure that the star was not variable, or that the brightening was not caused by some other effect.

Fortunately, any brightening caused by microlensing is expected to be characterized by a particular kind of light curve. The shape of this curve can be used to distinguish microlensing events from other types of variability. There are other checks too as the following question illustrates.

■ How would you expect a microlensing event to affect the colour of a star?

❑ Not at all. Light at all wavelengths should be affected in the same way, so microlensing will not affect the colour of the star.

Despite its difficulties, this technique has been successfully employed to search for MACHOs against the dense stellar background of the Galactic bulge and the two Magellanic Clouds, that are located just outside the Milky Way. Many dozens of microlensing events have been seen since the first detection in 1992. An example is shown in Figure 8.2.

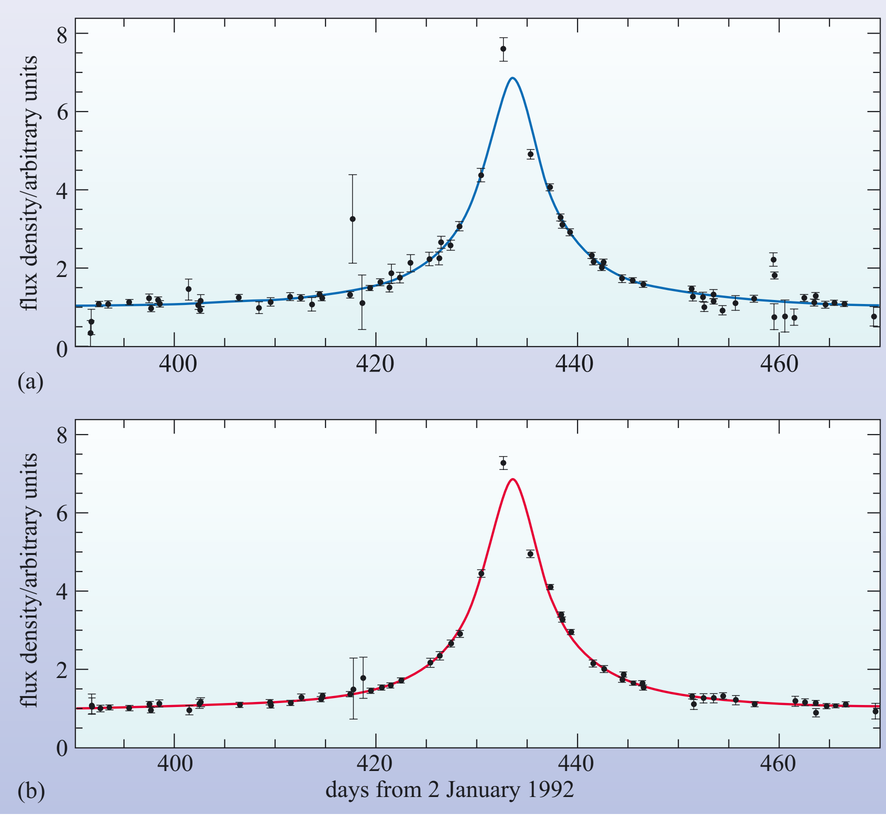


Figure 8.2 Gravitational microlensing. A star in the Large Magellanic Cloud brightens by a factor of 7 as an unseen dark object passes in front of it. The shape of the light curve is characteristic of microlensing and can be predicted from gravitational theory. As expected, the light curves are identical within experimental uncertainty in (a) blue and (b) red light. (Alcock *et al*., 1993)

The observations to date suggest that no more than about 20% of the mass of the Milky Way's dark-matter halo could be accounted for by MACHOs. So, although some of the dark matter in the halo may be MACHOs, most of it is not. In fact, this is not so surprising. As you have seen, there are very good theoretical reasons for believing that most of the dark matter cannot be made of MACHOs or any other kind of baryonic matter – we have to look elsewhere.

8.2.2 WIMPs

Weakly Interacting Massive Particles or WIMPs were introduced in Chapter 6 when we discussed the growth of structure under the influence of gravity in the early Universe. In that context, the justification for introducing the idea of a particle that responds only to gravity and the weak interaction was that density fluctuations composed of such particles would be able to grow prior to recombination. It should be stressed that at present no one knows what the WIMPs are, but several candidates have been proposed.

The original WIMP candidate, and the only one known to exist, is the neutrino. Neutrinos were plentiful in the early Universe, but ceased significant interactions with other forms of matter when the Universe was about 1 s old, just prior to the large-scale annihilation of electrons and positrons (see Section 6.3.7). As a result, the Universe should still be filled by a 'gas' of cosmic neutrinos, which should now have cooled to a temperature of roughly 2 K and a number density of about 10^8 per cubic metre. Individually, these neutrinos carry so little energy that they are currently undetectable, but collectively they might make a significant contribution to the density of dark matter.

Until recently it was thought that neutrinos had zero rest mass and always travelled at the speed of light. Such massless neutrinos would still contribute to the total density parameter Ω since this includes all kinds of energy, but the total contribution from massless neutrinos would be small. However, physicists have long recognized that there is no fundamental reason why neutrinos should not have mass, and neutrinos with mass, even a very small mass, might make a substantial contribution to the cosmic density. Recent measurements involving neutrinos from the Sun have indicated that the mass of a neutrino is roughly five million times smaller than the mass of the electron. Although this is an important result, such a tiny mass is not enough to let neutrinos account for a significant proportion of the dark matter. Neutrinos must make up some of the non-baryonic dark matter, but only, it seems, a tiny fraction amounting to about 0.3% of the total cosmic density.

Another serious objection to the neutrino is that because of its speed it could only be a candidate for *hot* dark matter. As was noted in Chapter 2, the hot dark matter scenario has fallen out of favour since it results in the formation of structure that is inconsistent with observations. For this reason, most cosmologists now reserve the term 'WIMP' for the proposed particles of *cold* dark matter.

Unlike the neutrino, all the candidates for cold WIMPs are hypothetical – no known particles fit the bill. There are several possibilities, but we shall consider only one, or rather only one class of candidates. This one class of candidate WIMPs is associated with a proposed new symmetry of nature known as **supersymmetry**. The notion of symmetry plays an important part in the standard model of elementary particles, since it implies various relationships between the fundamental particles and between the laws that govern them. Supersymmetry – first proposed

in the 1970s but still unconfirmed – would, if it existed, greatly extend the known symmetries of nature and imply the existence of many kinds of particles that have not yet been observed in nature.

Members of one class of supersymmetric particles, called **neutralinos**, have attracted particular attention from cosmologists. It is expected that there should be a stable neutralino with a relatively high mass – about 20 to 1000 times the mass of the proton. Like neutrinos, neutralinos are predicted to interact with other particles only through the weak interaction and through gravity. If the theory of supersymmetry is correct then some of the neutralinos created in the early moments of the big bang will still be present in the Universe today and these might be the WIMPs mainly responsible for cold dark matter. How could we find out whether neutralinos really exist?

There are two ways of going about this. One is to try to create a neutralino by simulating the high-energy conditions that were present in the big bang. This is not as dangerous as it sounds, since extreme conditions are routinely achieved in particle accelerators by making particles collide with each other at high energies. New particles are formed in the collision, and the hope is that, at the right energy, some of them will be neutralinos. Present-day accelerators are not powerful enough to achieve the required energies, but the new Large Hadron Collider being built at the European Laboratory for Particle Physics (CERN) near Geneva, may be able to do it (Figure 8.3). Early experiments with less powerful colliders have failed to create a neutralino but they have shown that, if it exists, the neutralino mass must be at least 30 times the mass of the proton.

Figure 8.3 The Large Hadron Collider at CERN near Geneva will occupy a circular tunnel 27 km in circumference, which currently houses the Large Electron–Positron Collider. When operational it may be possible to create neutralinos, the candidate particles of cold dark matter. (CERN)

Another approach is to try to detect cosmic WIMPs directly. You saw in Chapter 6 how particles of high mass could be created from radiation in the very earliest moments of the big bang when temperatures were extremely high. It is possible to estimate how many neutralinos might have been formed and so gauge how likely we are to detect them.

Although dark matter in the Milky Way is believed to form a widely dispersed dark-matter halo, the visible parts of the Galaxy lie within this halo. Thus the dark matter particles are expected to permeate the entire Galaxy, and we may expect to find neutralinos in our neighbourhood. They interact only weakly with ordinary matter, so they will pass freely through the Earth. The predicted number density of these particles is such that over ten million of them will pass through your head as you read this sentence.

This opens up the attractive possibility of detecting neutralinos in a laboratory experiment. Many such experiments are under way. They all work on the principle that if presented with a suitable target, a small proportion of the neutralinos passing through the Earth might interact with that target to produce a tiny flash of light or a minute rise in temperature.

Several materials are being used as targets but sodium iodide, in the form of large crystals up to 100 kg in mass, is a common one. The sodium and iodine atoms will recoil when hit head-on by a neutralino, and the recoil energy will be released as a flash of light that can be detected. Estimates suggest that a 10 kg crystal could be used to detect one neutralino a day. The biggest problem is to distinguish between genuine neutralinos and various kinds of background radiation. One series of experiments, by the UK Dark Matter Collaboration, is being conducted 1100 metres below ground in a potash mine where the detectors are shielded from cosmic rays (Figure 8.4). Other events can cause flashes of light, too, and the physicists work on the principle that when all other possible events have been eliminated then whatever remains must be a neutralino. At the time of writing, none of the various experiments around the world have unambiguously detected a neutralino. However, detector sensitivity is improving all the time, and if a future experiment does give a positive detection of a dark matter particle it will mark a major advance in cosmology.

QUESTION 8.2

Table 8.1 presents a number of candidates for dark matter. Classify each candidate by placing ticks in the appropriate columns.

Table 8.1 For use with Question 8.2.

Dark matter candidate	Baryonic	Non-baryonic	MACHO	WIMP	Cold	Hot
brown dwarfs						
neutrinos						
neutron stars						
black holes						
neutralinos						

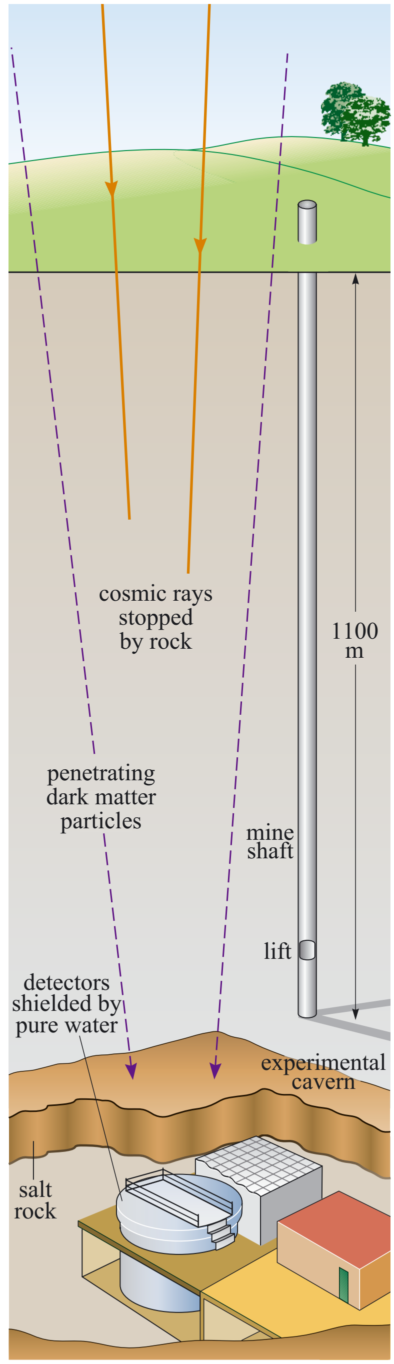


Figure 8.4 The UK Dark Matter Collaboration's WIMP detector is buried 1100 metres below ground at the Boulby Mine in the North York Moors National Park (in northern England). The deep location of the detector provides some shielding from cosmic rays, which are a source of 'noise' for dark matter experiments.

8.3 The nature of dark energy

While the prospect of a Universe filled with exotic new matter has itself challenged our understanding of the physical world, an even more startling cosmological discovery has come to light in the last few years. Until the late 1990s cosmologists took it for granted that the expansion of the Universe was slowing down under the mutual gravitational attraction of the matter within it. The main question was whether the expansion would eventually stop and go into reverse, or else carry on forever. Things are no longer so simple. As we saw in Chapter 7, studies of remote Type Ia supernovae imply that cosmic expansion is not slowing down but speeding up, and recent measurements of CMB anisotropies have given support to this finding. The accelerating expansion can be accounted for by attributing about 73% of the energy density of the Universe to a so-called dark energy, the effect of which is to oppose the deceleration of cosmic expansion. The presence of dark energy allows the Universe to maintain the critical density, even though the density of matter is low and decreases with time.

■ What is the connection between dark energy and dark matter?

❑ None that we know of! Although the names are similar, they refer to completely different phenomena. In particular, dark energy is *not* the energy equivalent of dark matter.

The dark energy acts to oppose the gravitational attraction of the matter in the Universe. But what does that mean physically? In general relativity, gravity is a consequence of space–time curvature, and that is determined by the distribution of energy and momentum, not just by the presence of massive bodies. Even the pressure in a fluid influences its gravitational effect since, as the fluid expands or contracts, the pressure will affect the internal energy of the fluid, and this energy will influence the curvature of space. (This is one way in which general relativity differs from Newtonian gravity.) A normal fluid uniformly filling the Universe, as envisaged in the Friedmann–Robertson–Walker models, would exert a positive pressure at every point, and this, like the density of the fluid, would have the gravitational effect of decelerating the cosmic expansion. The accelerating effect of dark energy indicates that it, in contrast to a normal fluid, is a source of *negative* pressure. The *density* of dark energy actually tends to retard cosmic expansion, but the negative pressure more than compensates for this, so the overall effect of dark energy is to accelerate the expansion. The negative pressure, or rather the associated gravitational effect, can be thought of as driving the cosmic acceleration. (Incidentally, if you think that *positive* pressure is normally responsible for pushing things apart, you are probably thinking of the effect of *differences* in pressure between one region and another. FRW cosmology is concerned with uniform universes, where there are no pressure differences, but where uniform pressure, like uniform density, does have a gravitational effect.)

Negative pressure may be unfamiliar but it is not unphysical. A phenomenon known as the **Casimir effect** (see Figure 8.5) shows that the presence of two narrowly separated, parallel metal plates modifies the electrical properties of the vacuum between them, producing a negative pressure in that region. In this case there *is* a pressure difference, and it creates an effective attraction between the plates that can be demonstrated and measured experimentally. This is not a gravitational effect, but it is a demonstration of negative pressure. Don't worry if you find the idea of

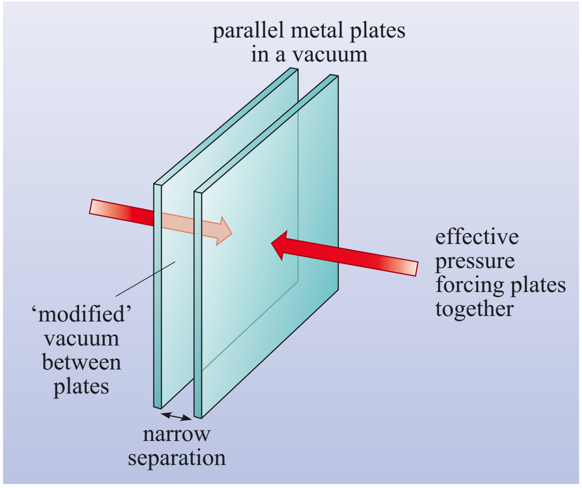


Figure 8.5 The Casimir effect. Two narrowly separated, uncharged conducting plates, located in a vacuum will be attracted towards one another. The attraction arises from the influence that the plates have on the region between them and the negative pressure that this produces in that region.

negative pressure hard to grasp. Accept for now that a uniform distribution of dark energy with negative pressure will have a repulsive gravitational effect, and such a distribution of dark energy will tend to accelerate the Hubble expansion.

What could this mysterious (and, let's face it, rather weird) dark energy possibly be? This is our Problem 2, and cosmologists have come up with three plausible answers.

The first possibility is that the dark energy simply represents the effect of a cosmological constant Λ, and has no deeper explanation. Unlike matter and radiation, the energy associated with the cosmological constant would not be diluted by the expansion of the Universe. It would stay the same, exerting a constant negative pressure throughout the expansion of the Universe. The value of the constant, Λ, would then be a new constant of nature, much like the gravitational constant, G. We would not be able to explain why it had the value it had, any more than we can explain why G has the value it has. It's just one of those things. Many cosmologists are unhappy about simply attributing the dark energy to the cosmological constant. Because its value is arbitrary and can be chosen to fit any acceleration, it seems too much of a 'just so' explanation.

The second possibility comes not from general relativity but from quantum physics. One of the central features of quantum theory is **Heisenberg's uncertainty principle**, which can be expressed in several ways, including the simple formula

$$\Delta E \Delta t > h/2\pi \qquad (8.1)$$

where h is the Planck constant. Its usual interpretation is that we cannot know both the precise energy of a particle and the precise time we measure that energy. If we wish to know the energy of a particle to an uncertainty ΔE, then the time we take to measure it must be at least Δt. We cannot know both the energy and the time to greater precision.

The implications of this are profound and not a little disturbing when we apply it to empty space. Let's do a thought experiment. Suppose we take a small box, as small as we wish, and clear it of all particles. We also shield it from the outside world to make sure no fields are present within it. It's as empty as we can get it.

- What would be the energy density inside this box?
- Common sense tells is that if the box is empty the energy density inside must be zero. But the uncertainty principle tells us otherwise!

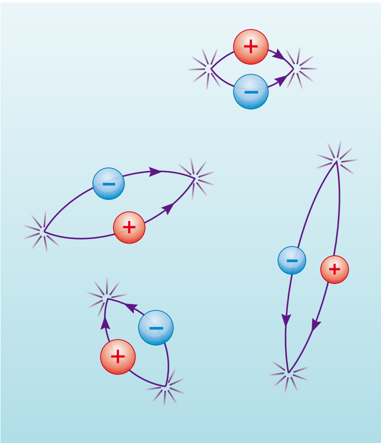


Figure 8.6 Virtual particles. According to quantum physics pairs of particles and antiparticles are continually being created and destroyed in empty space. The more massive the particle, the shorter its life.

If space were devoid of particles we would be able to say 'the energy in this part of space at this time is zero'. But we would then know the energy exactly and that would violate the uncertainty principle! The uncertainty principle forces us to recognize that even in 'empty' space, particles of energy ΔE could exist for a time Δt. It also implies that more massive particles (with greater mass energy) will be shorter lived than less massive particles. These **virtual particles**, which always appear as pairs of particles and antiparticles, are created and destroyed before they can be observed. (The process is similar to the process of pair production that was discussed in Chapter 6.) Space, far from being empty, is teeming with particles continually popping in and out of existence (Figure 8.6). The collective energy of these particles is known as **vacuum energy**.

There is no doubt that vacuum energy exists. It provides the explanation of the Casimir effect. The presence of the parallel conducting plates limits the kinds of virtual particles that can form in the region between the plates, and it is this that causes that region of 'empty space' to have different properties from the surrounding 'empty space'. The fact that vacuum energy is a property of the vacuum itself ensures that it will not be 'diluted' by the expansion of the Universe, and that it will have the required negative pressure.

In fact, the vacuum energy would have the same effect as a cosmological constant. Nonetheless, a distinction should be drawn between the cosmological constant as an irreducible term in the field equations of general relativity, and the quantum physical vacuum energy as a particular contribution to the cosmic distribution of energy and momentum. They may have similar effects, and either may account for the dark energy, but strictly speaking, they have a different origin.

This similarity has led to a lot of confusion, with some authors using the terms dark energy, vacuum energy and cosmological constant interchangeably. But it also means that the behaviour of possible universes, as described by the equations of Chapter 5 and 6, is just the same whether Λ represents Einstein's cosmological constant or a quantum vacuum energy with the same properties.

There is one important difference, though. Since vacuum energy is a consequence of quantum physics, it should be possible to calculate its density from first principles, to see how it compares with the measured density of the dark energy. This is something that cannot be done for the cosmological constant.

Can we calculate the density of the vacuum energy? Yes we can, though with some difficulty, and it comes out to about a factor of 10^{120} – 120 orders of magnitude – higher than the measured density of dark energy! This seems bizarre. Can there really be so much energy in empty space? Wouldn't we notice it? The repulsive gravitational effect of so much vacuum energy would be so great that the Universe would expand explosively (much like inflation in fact, where the vacuum energy is also implicated, as you saw in Chapter 6). If the vacuum energy is as high as it seems it's a real puzzle why it does not make itself felt.

Despite the huge discrepancy between the observed and calculated energy densities (Nobel laureate Steven Weinberg has called it 'the worst failure of an order-of-magnitude estimate in the history of science'), theoretical physicists seek solace in the knowledge that our understanding of particle physics is not complete. Until it is, they cannot allow for all the possible particles that may appear in the vacuum. In some theories particle energies may cancel each other out, so physicists hope that when a more complete theory is available the vacuum energy may yet turn out to account for the dark energy, but there is no proof of that at present.

The third candidate for dark energy is called **quintessence**. The name alludes to the fifth element of the ancient Greeks (after earth, air, fire and water), which was supposed to constitute the heavenly bodies. In its modern cosmological usage, quintessence can be thought of as an exotic form of matter. However, to account successfully for the observed properties of dark energy, quintessence would have to be such an odd form of matter that it actually makes more sense to think of it as a distribution of energy that fills the Universe. This makes it sound like vacuum energy, but unlike vacuum energy, quintessence is assumed to vary in time and space. Vacuum energy is a very specific form of energy whereas quintessence encompasses many possibilities. Quintessence would have to exert negative pressure to cause the observed acceleration in cosmic expansion, but its properties might be such that in the early Universe it behaved in a way similar to the energy of ordinary radiation and, though dominant at present, its influence might decline in the future allowing the cosmic acceleration to cease and even permitting the eventual deceleration of the Universe. Changes in the rate of cosmic acceleration should make it possible to distinguish the effects of quintessence from those of vacuum energy or the cosmological constant. Such measurements are not conclusive at present, but they may become so in future with improved observations of supernovae and CMB anisotropies.

Figure 8.7 summarizes the three main candidates for dark energy. So you thought space was empty? US physicist Hans Christian von Baeyer put it rather well. 'Space,' he said, 'is empty of matter but filled with surprises.'

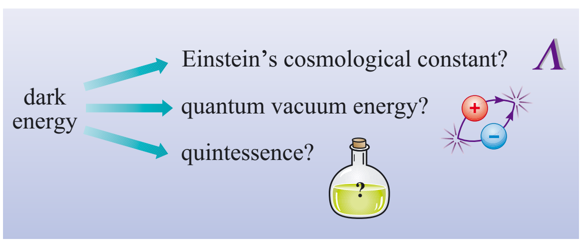


Figure 8.7 The three main candidates for dark energy.

8.4 The horizon and flatness problems

Two observed properties of the Universe that should be accounted for by cosmological models relate to its uniformity and the flatness of its three-dimensional spatial geometry. Neither of these properties is a natural outcome of the hot big bang model. Consequently the issues of explaining the uniformity and spatial flatness of the Universe have been labelled as 'problems' called, respectively, the horizon and flatness problems.

We start by considering the problem of explaining the uniformity of the Universe (Problem 3 in our list).

8.4.1 The horizon problem

You first came across the horizon problem in Chapter 6, when we considered the uniformity of the cosmic background radiation. This problem arises from the observation that the temperature of the cosmic background radiation is uniform to a few parts in 10^5 across the sky, yet points on the sky more than about two degrees apart are separated by a distance that is greater than the *horizon distance* at the time of last scattering (Figure 8.8). Remember that the horizon distance at a given time represents the maximum distance that a physical signal could propagate through space in the time elapsed since the very first instant of the big bang.

Although we have concentrated on the horizon distance at the time of last scattering, it is possible to consider the horizon distance and uniformity of the Universe at later times, and to arrive at a similar conclusion about the existence of the horizon problem. The advantage of discussing horizon distances in terms of the last-scattering surface is that it is readily observable through observations of the CMB. Thus, according to the standard account of the big bang, there is no reason to expect the CMB to be uniform on scales greater than about 2°.

One solution to the horizon problem is that, despite the arguments we have advanced to suggest that regions of the last-scattering surface could not have been in communication with each other, nevertheless, they *have* colluded together. Somehow or other, in the very early stages of the Universe's development, they *were* in contact with each other and reached a state of thermal equilibrium. What is required then is that these regions were somehow separated. A key idea here is that such a separation would create large-scale uniformity by the expansion *of space* rather than the propagation of a physical signal *through space*. The expansion would have to be by an enormous factor so that two regions that were once in contact, would later appear to be separated by many horizon distances.

- ■ What process in the early Universe could have had this effect?
- ❑ The process of inflation – a brief period of very rapid expansion could have caused regions that were once in contact to become separated by more than the horizon distance. (Inflation was discussed in Chapter 6.)

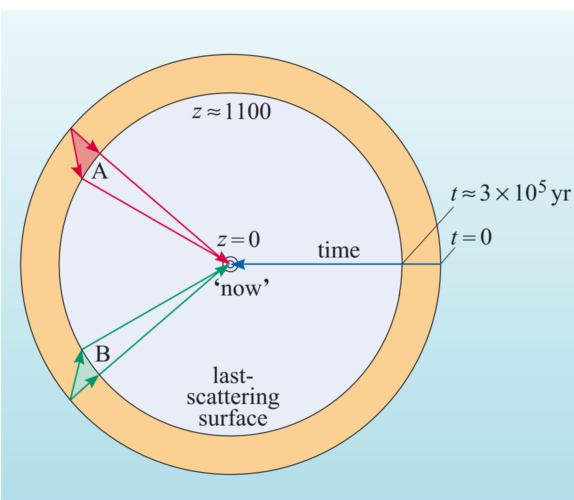


Figure 8.8 Regions A and B individually have an extent equal to the horizon distance at the time of last scattering. Both A and B subtend an angle of about 2° on the last-scattering surface (note that these angles are exaggerated in this diagram). However, A and B are separated by an angle greater than 2° and so lie outside each other's horizon at the time of last scattering, and hence no physical communication could have occurred between these two regions. The fact that the CMB appears to have the same temperature at A and B, despite this apparent lack of communication, is an example of the horizon problem.

To recap, the key idea of inflation is that close to the time that the grand unified era came to an end, the Universe underwent a brief but rapid phase of exponential expansion. The net result was that by the time that inflation was over the scale factor had increased by a huge amount. As noted in Chapter 6, the growth in the scale factor during inflation is very uncertain – but it is thought that an increase by a factor of as much as 10^{50} or more could have occurred during this process. Thus a region of the Universe, having first had a chance to become homogeneous and to reach thermal equilibrium before inflation, might have expanded to a size that is much greater than the horizon distance.

So inflation can solve the horizon problem by enlarging and sweeping apart regions of the Universe that had already become homogenized and leaving them separated by vast distances. When the cosmic background radiation became decoupled from matter (when the age of the Universe was about 380 000 years), it was necessarily isotropic because it was emitted by matter that was already homogeneous.

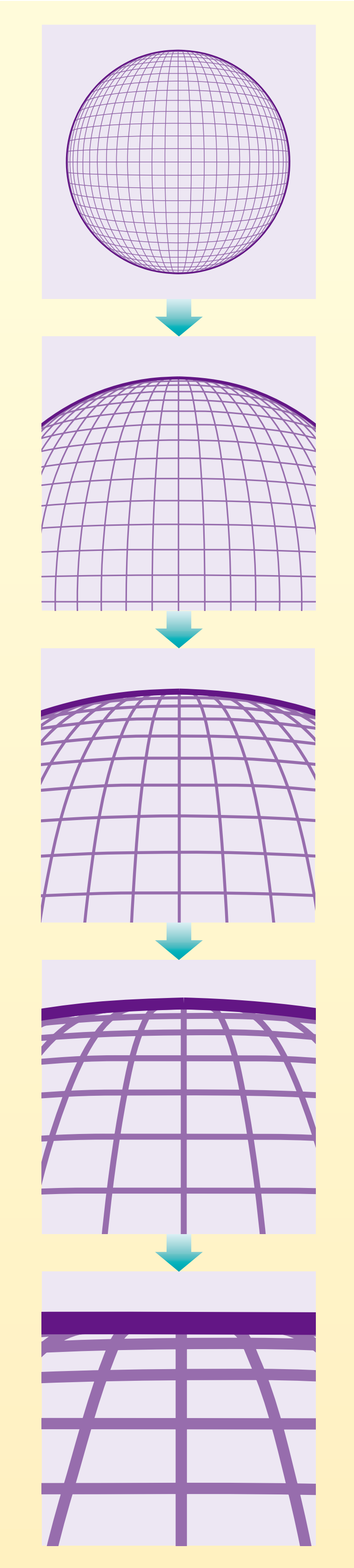
It is important to appreciate that inflation should be considered as an addition to the standard hot big bang model. Whereas the predictions of the standard hot big bang model from times of about 10^{-9} s onwards are widely accepted and stand up to observational scrutiny, the inflationary hypothesis is far more speculative. Indeed, as we noted in Chapter 6, the mechanism that drives the process of inflation is essentially unknown. The appealing aspect of inflation is that it offers a single solution to several cosmological problems; not only the horizon problem that we have just considered, but also the problems of structure formation and spatial flatness. It is to the latter problem that we now turn.

8.4.2 The flatness problem

The next outstanding problem on our list, Problem 4, relates to the observation that the average density of the Universe is almost equal to the critical density, i.e. $\Omega = 1$. In the context of the Friedmann–Robertson–Walker models this causes the curvature parameter, k, to be zero (see Question 5.9) and implies that three-dimensional space will have a flat geometry. Why should this be a problem?

In order for the total density parameter, $\Omega(t)$, to be close to 1 today, it had to be even closer to 1 in the past. This is because the extent to which $\Omega(t)$ differs from 1 is predicted to grow with time. If $\Omega(t)$ were *exactly* equal to 1 today, at time t_0, then at any earlier time, it would also have been exactly equal to 1. If, however, the current value of the density parameter turned out to be somewhat less than 1, $\Omega(t_0) = 0.90$ say, what would its value have been at some earlier time when the age of the Universe was only a fraction of t_0? The answer depends on the details of the FRW model used to represent the Universe, but in one case, for example, the difference grows at a rate proportional to $t^{2/3}$. In this particular case, a current difference of 0.1 implies that when the Universe was a thirtieth of its present age the difference between $\Omega(t)$ and 1 would have been smaller by a factor $30^{2/3} \approx 10$. Thus if $\Omega(t_0) = 0.90$, then $\Omega(t_0/30) \approx 0.99$. The fact that the Universe is now about 10^{17} seconds old, and has a density parameter that is still close to 1 means that at very early times, $t < 10^{-6}$ s say, $\Omega(t)$ must have been *extremely* close to 1. Explaining why $\Omega(t)$ should be so close to 1 at very early times is the crux of the flatness problem.

Now, of course, it is possible that the total density of the Universe just happens to have the critical value, in which case $k = 0$ and $\Omega(t)$ will always be equal to one.



However, this would be another of those 'just so' explanations that, though possible, are never much trusted by cosmologists. Their preference is always for 'mechanisms' or 'processes' that force the cosmological parameters to take on their observed values. One of the motivations for proposing that there might have been an era of inflation in the very early Universe is that this can solve the flatness problem just as neatly as it solves the horizon problem.

The most direct result of inflation is an enormous increase in the cosmic scale factor $R(t)$, perhaps by a factor of 10^{50} or more. As suggested by Figure 8.9, this will have the effect of reducing the curvature of space, which depends on the quantity k/R^2. A sufficient amount of inflation will result in such a small value of k/R^2 that the effective value of k is zero, and space is geometrically 'flat', irrespective of the true value of k prior to inflation.

Now a geometrical argument of this kind might seem quite convincing at first sight, but you might still wonder how such an argument can have any bearing on the density of the Universe. Figure 8.9 might suggest that, in effect, $k = 0$, but how can it account for the fact that $\Omega(t)$ is correspondingly close to 1? To understand this you have to recognize that the link between the effective curvature and the density comes directly from the Friedmann equation and remains true no matter what the source of the cosmic energy density may be. Thus, near the end of inflation, when space is effectively flat, the Universe might have become quite cold and most of its energy might take the form of some exotic kind of vacuum energy, but the density of that vacuum energy will be just what is required to produce an effectively flat geometry. The usual assumption is that as the Universe ceases to inflate, some of this vacuum energy is converted into more conventional forms of matter and radiation, and that the Universe is reheated to a temperature similar to that it would have had in the absence of inflation. This has the interesting effect of causing all the matter and radiation in the observable Universe to be a direct consequence of inflation (any pre-existing matter or radiation will have been so diluted by inflation as to be unobservable), but it will not alter the fact that when all forms of matter and radiation are taken into account, as well as any dark energy, the total density of the Universe will be very close to the critical density and $\Omega(t)$ will be correspondingly close to 1.

Though inflation leads in a natural way to a Universe that is close to having to a critical density, it does not convert the Universe into one that has *exactly* the critical density. If the density before the onset of inflation was greater (or less) than critical, it will still be greater (or less) afterwards, though only barely so.

■ Supposing that prior to inflation the density were *less* than critical, what would be the analogue of Figure 8.9 for a Universe undergoing inflation?

❑ It would be a saddle-shaped rubber sheet which, after inflation, had been almost completely flattened over the tiny part constituting the 'observable' Universe.

Figure 8.9 A spherical balloon analogue for illustrating how the inflation of the Universe at an early epoch would have resulted in a flat spatial geometry for the observable Universe regardless of the curvature prior to inflation.

As you have seen in Chapter 7, the best current measurement of the cosmic density parameter is 1.02 ± 0.02, showing that the Universe is indeed very close to the critical density and very close to having a flat geometry. You have also seen that the contribution from the dark energy is about 73% of the critical density with matter at 27%. Hence the observation that two apparently unconnected components of the Universe, matter and dark energy, add up to the critical density is simply explained by inflation. Otherwise this coincidence seems very hard to understand.

At a stroke, the inflation idea solves both the horizon *and* flatness problems that afflict the standard model of the big bang. We have solved Problems 3 and 4! Indeed, it might even be that inflation holds the key to understanding the slight inhomogeneities that do occur – those responsible for triggering the formation of galaxies and those that manifest themselves as the ripples in the microwave background radiation. We shall learn more about this in the next section.

8.5 The origin of structure

In Chapter 2 you were introduced to the idea that galaxies formed from primordial fluctuations in the density of matter in the early Universe. Later, in Chapter 4, you saw how astronomers study the large-scale distribution of galaxies to test theories of how galaxies formed and infer something about the nature of the early fluctuations. Then, in Chapters 6 and 7 you learned how these fluctuations left their imprint on the cosmic background radiation.

Now it is time to address our Problem 5: where did those primordial fluctuations come from? In a sense this is the inverse of the uniformity problem. Earlier we asked why is the Universe so uniform, and now we ask why it is not perfectly uniform. From where did all this diverse structure come?

Inflation enlarges a tiny region of space by a huge factor. Solving the horizon problem may require space to expand by many orders of magnitude. This implies that current cosmological scales, tens to hundreds of megaparsecs, corresponded to subatomic scales during inflation. This intense magnification means that large-scale structures in the present-day Universe may have been rooted in subatomic irregularities in the pre-inflationary Universe. But where could such irregularities have come from in the first place?

As you have seen in the discussion of vacuum energy, empty space is seething with virtual particles that flit in and out of existence as a consequence of the Heisenberg uncertainty principle. Similar **quantum fluctuations** occurring during inflation would have been enlarged by the very rapid expansion of the Universe and could have been the cause of macroscopic variations in density. These, after much subsequent evolution might have been the source of the large-scale structure we see around us today.

Although there is no agreement about exactly what caused inflation, there are a number of proposed models that make quite detailed predictions about the quantum fluctuations that might have occurred. Generally speaking these fluctuations would have been caught up in inflation and would have grown in size until they exceeded the horizon scale of the then-observable Universe (Figure 8.10). Once this happened, information could no longer travel from one side of a fluctuation to the other, so the fluctuation would have been unable to smooth itself out and would have become ‘frozen in’, expanding along with space as the scale factor $R(t)$ continued to grow.

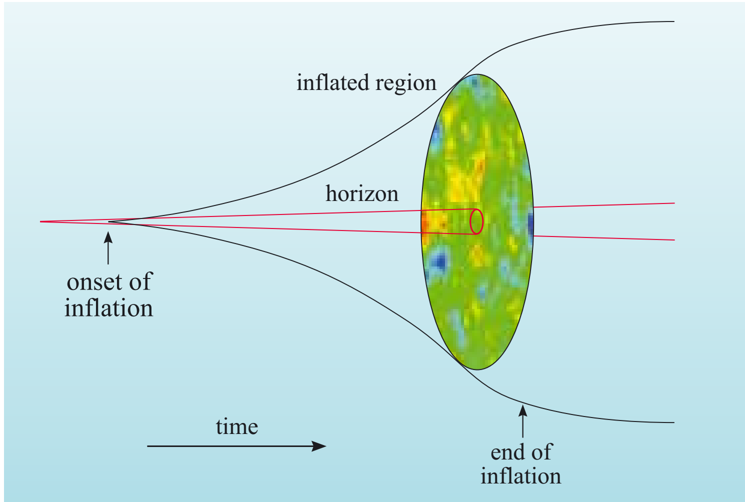


Figure 8.10 A schematic illustration of the expansion of a small region of the Universe due to inflation. Quantum fluctuations within this region are expanded beyond the horizon distance, at which point they become 'frozen in' and form the primordial density fluctuations from which subsequent structure in the Universe develops.

As earlier fluctuations were stretched beyond the horizon scale, newer quantum fluctuations were generated, creating more small-scale density variations that were stretched in their turn. The precise outcome varies from one model of inflation to another, but the general result is a range of density fluctuations with roughly the same 'strength' on a wide range of size scales, i.e. variations in density that are much the same on large scales and on small scales.

Following inflation, the Universe continued to expand, but the rate of expansion was so much reduced that the growing cosmic horizon would have encompassed more and more of the matter in the Universe as time passed. As the horizon expanded at the speed of light, material that had been swept over the horizon during inflation re-entered the horizon bringing the frozen-in density variations with it. The last density fluctuations to be inflated beyond the horizon would have been the first to re-enter, but these would have been followed by other density variations on larger and larger scales. Once back inside the horizon, the variations in density ceased to be 'frozen-in' and were able to become stronger or weaker depending on the prevailing conditions at the time they re-entered the horizon. It has been known for some time that the pattern of density variations predicted by inflation (with the same amplitude on all scales) is just what is required to give rise to the observed range of superclusters, clusters, etc. Hence inflation provides a possible explanation for the origin of all of the structure we see in the Universe – it might all have come from quantum fluctuations. Intriguingly, according to the inflationary hypothesis, the largest macroscopic structures might have had their origin in the microscopic quantum world.

8.6 The matter of antimatter

Inflation can explain many things, but it cannot account for the fact – Problem 6 – that the Universe contains far more matter than antimatter. Why should this be so?

In the account of inflation given in Chapter 6, it was noted that the end of inflation would have been accompanied by the release of a vast amount of energy in the form of particle–antiparticle pairs. At a later stage the antiparticles annihilated with the particles, so by now we might expect the Universe to be devoid of matter but full of radiation. In numerical terms this is very nearly the case: there are about a

billion photons (mainly in the cosmic microwave background) for each proton or electron in the Universe. Even so, the fact that there are any matter particles at all, implies that rather more matter than antimatter must have been formed.

Another way of stating this problem is in terms of the baryon number of the Universe. If, for every baryon in the Universe, there was a corresponding antibaryon, the baryon number of the Universe would be zero. However, we know that in the real Universe there is a surplus of baryons over antibaryons – so the baryon number of the Universe is a positive number. The conservation of baryon number in nuclear reactions is an important principle of physics, and we might naively expect that because the baryon number of the Universe currently has a positive value, this must always have been the case. However, physicists are reluctant to accept that the Universe must have started out with a positive baryon number. It seems more natural for the baryon number of the Universe to have originally been zero – in much the same way that the net electric charge of the Universe is zero. So if the Universe started out with a baryon number of zero – how did it reach its present non-zero value? One possible answer is that there may have been an era in the history of the Universe when baryon number was *not* conserved, i.e. reactions may have occurred which violate the principle of conservation of baryon number. Such reactions could have caused the baryon number of the Universe to change from an initial value of zero to the positive value we observe in the present-day Universe.

In Chapter 6 you saw that the unification of the strong and electroweak forces (grand unification) is believed to occur at energies of around 10^{15} GeV. The speculative grand unified theories that describe reactions at these enormous energies predict that baryon number need not always be conserved. Particle interactions at such energies could therefore give rise to the slight excess of matter that we see in the present-day Universe. Note that such interactions must have occurred *after* the process of inflation. According to the inflationary model, the vast majority of particles in the Universe were created from the energy released at the end of inflation. The imbalance between matter and antimatter could only have developed after these particles were created, and so inflation must have occurred at, or before, the end of the grand unified era.

Reactions in which baryon number is not conserved occur at such high energies that they are far beyond anything we can reproduce in laboratories today. However, it might still be possible to find experimental evidence in support of grand unification theories. The fundamental processes that allowed an excess of baryons to form in the very early Universe might also, very rarely, allow protons to decay in the present-day Universe. So far, no proton has ever been observed to decay, despite many attempts to detect such a process. However, if the proton had a mean life of the order of 10^{33} years, this failure could be understood and there would still exist the possibility of observing proton decay in the future.

Before we leave this subject, we need to address an assumption we have made, namely that there are no significant amounts of antimatter in today's Universe. How do we know this? If matter and antimatter could somehow have become segregated in the early Universe, then there may be regions of space in which antimatter dominates. Antiprotons, antineutrons and antielectrons (i.e. positrons) could have come together to form antiatoms which in turn could have formed antimolecules and eventually antistars, antiplanets and antipeople. Are there intelligent creatures out there made of antimatter?

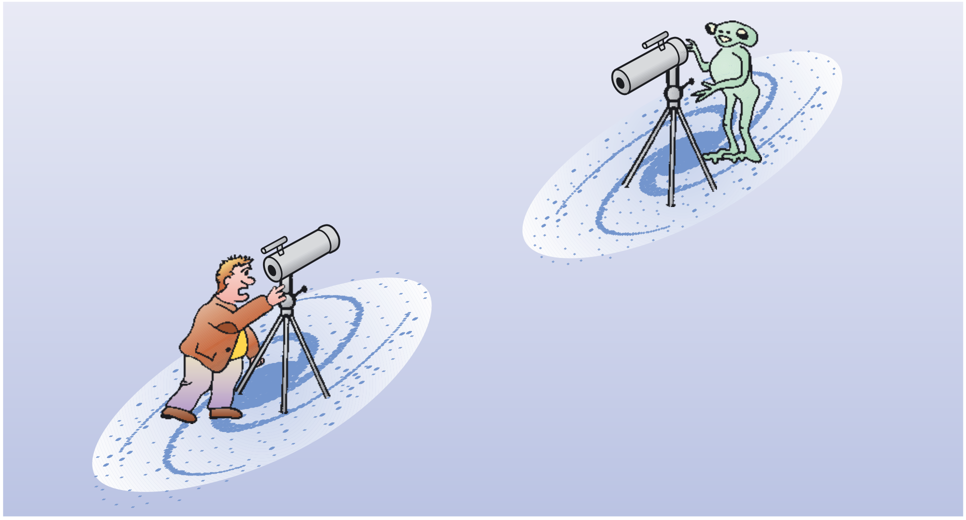


Figure 8.11 How could we tell whether a distant galaxy was made of antimatter?

QUESTION 8.3

Suppose you suspected that a newly discovered galaxy was made of antimatter (Figure 8.11).

(a) Could you tell from its emitted radiation whether the galaxy was made of matter or antimatter?

(b) Given that matter and antimatter will annihilate each other to form γ-rays, are there any other observations you could make?

8.7 Towards $t = 0$

According to the classical Friedmann–Robertson–Walker models, the Universe started expanding from a condition in which the scale factor was zero, implying a state of infinite density, often referred to as the **initial singularity**. How can we talk about the physics of something in that state? What actually happened at $t = 0$? This is our Problem 7, and it is a difficult one. If we naively extrapolate the big bang model back towards $t = 0$, many of its physical properties (the energy density of matter and radiation, the pressure and temperature, and the curvature of space–time) approach infinity, i.e. they diverge. When a model predicts infinite values we can take it as a warning that we have probably pushed the model beyond its limits of validity.

This should not come as a great surprise. You saw in Chapter 6 that quantum physics sets a natural limit – represented by the Planck time, about 10^{-43} seconds – on the earliest moment at which we can have any confidence in the big bang model. As we approach the Planck time quantum effects become as important as general relativistic effects and behaviour cannot be understood within the framework of existing physical theory. Neither general relativity nor conventional quantum physics are much help to us here. So before we can think about $t = 0$ we need a new theory that will take us back beyond the Planck time. First, however, we consider the possible cause of inflation.

8.7.1 The origin of inflation

We would like to know how inflation began. There is no agreement about the fundamental physics underlying the inflationary hypothesis, unlike, say the big bang model, which is based on general relativity. Inflation implies a wealth of physics beyond what we already understand. One of the major aims of current research involves trying to link inflation to known particle physics.

Without a fundamental theory backing it, the question of how an inflationary period may arise in the Universe is open to speculation. Inflating even a small initial patch to form the observable Universe we see today puts restrictions on what came before. For instance, any initial patch must have been sufficiently uniform on average to allow inflation. How did this arise? Were these conditions predetermined – or did they arise by chance? As one possible solution to this, the cosmologist Andrei Linde has suggested a model called **chaotic inflation**, where the initial condition of the Universe is random. In this model the Universe is much larger than we can see, and is partitioned into *domains* with differing laws of physics (see Figure 8.12). We live in a region where the right conditions for inflation have arisen by chance. This region has inflated and has provided the right conditions for life, but other domains may not be capable of this.

A problem with this suggestion is that we cannot probe these other domains unless they impinge on our domain. This model is therefore not testable. One day, inflation may be found to be part of testable theory that explains how the right conditions could arise. At present however, it is not clear if this will ever be the case.

Inflationary theory has dominated thinking about the very early Universe since the 1980s. Its ability to explain a range of diverse features of the Universe with a small number of assumptions makes it an attractive model. The main ‘predictions’ of inflation – a spatially flat Universe that is highly but not perfectly uniform, and a specific spectrum of density variations – appear to have been validated by observation. This success has led cosmologists to take seriously the key elements of the inflationary scenario, but it has also left them with a strong awareness of the need for further tests and an interest in searching for alternative theories that might be equally successful. The next two subsections briefly discuss two of these alternatives.

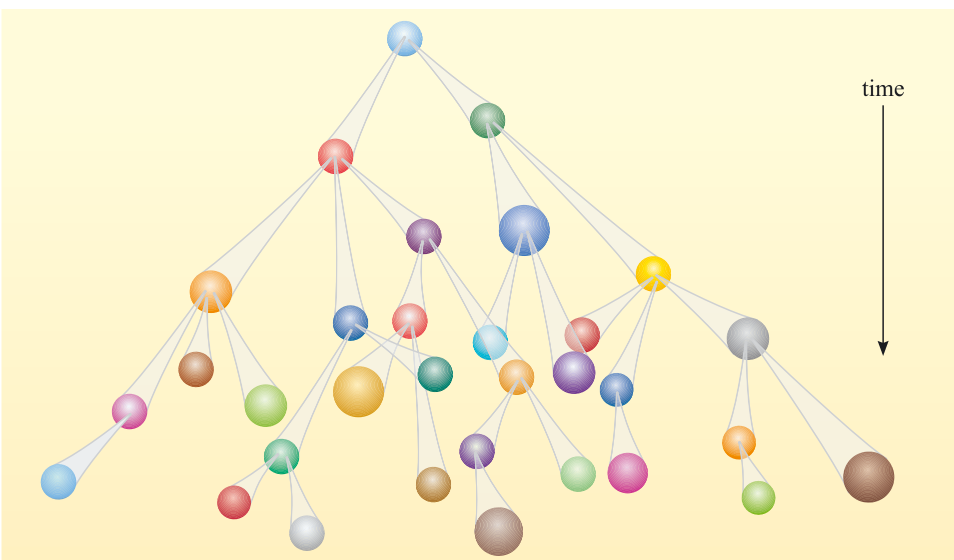


Figure 8.12 Chaotic inflation suggests that the Universe may consist of numerous inflationary domains each with different laws of physics (different colours in this diagram). Our observable Universe is just a tiny part of one of these domains.

8.7.2 Quantum cosmology

Quantum cosmology is used to describe one particular attempt to describe what went on before the Planck time. It employs a particular (and highly speculative) approach to the unification of quantum physics and general relativity that was pioneered by John Wheeler and Bryce DeWitt in the 1960s. Unfortunately, their theory is difficult to interpret as it does not refer directly to evolution in time, but rather to other properties of the Universe, such as its size. Remarkably, the Wheeler–DeWitt theory can be used to calculate the probability that the Universe came into being from nothing, though doing so involves making a number of assumptions. James Hartle and Stephen Hawking attempted this in 1983, basing their thinking on the assumption that the Universe should have no boundary in time or space. Others, using different assumptions have found a different answer and it is not clear which, if any, is correct. However it is interesting that one can make progress by assuming that the laws of physics are somehow 'beyond' the Universe itself and can describe its creation. This is in distinct contrast to the philosophy of chaotic inflation, where it is assumed that the laws of physics are very much 'within' the Universe, and may vary from place to place.

8.7.3 M-theory

Over the past few years a new line of research has opened up in cosmology. **M-theory** (which subsumes the earlier ideas of *superstring theory*) tries to unify not only general relativity and quantum theory, but also all the known forces in the Universe. It is widely regarded as the best candidate so far for a 'theory of everything.' In M-theory the fundamental objects are not particles, as usually assumed, but strings or even sheets. These strings and sheets can vibrate, and the different vibrations describe the different particles and their masses. One of the attractive features of these models is that they predict a massless particle that looks like the graviton, the quantum particle of the gravitational field. Hence M-theory 'predicts' gravity and all of the results of general relativity.

One of the stranger features of M-theory is that it predicts more than four dimensions. In fact it requires 11 dimensions to work at all. Usually it is assumed that the extra dimensions are 'rolled up' and hidden away to leave the familiar three dimensions of space and one of time – the extra dimensions would only become apparent at very high energies. However it seems that some of the extra dimensions need not be rolled up. One line of development regards our Universe as residing on a sheet (known as a 'brane' since it is derived from the term membrane) that itself resides in a five-dimensional space called the bulk. This picture has allowed cosmologists to develop new models of the early Universe, including one that does not require inflation to have occurred at all. According to this **ekpyrotic model** (the name means 'out of fire'), a key event in cosmic history was a collision between the brane on which our Universe resides and some other 'parallel' brane (Figure 8.13). This collision would have been the cause of the fluctuations that led to large-scale structures being formed. In an extension of this model, the sheets collide over and again, allowing an infinite number of past big bangs in a 'cyclic' universe. Just before the collision, the forces on the sheets cause an accelerated expansion, which would look like the currently observed acceleration, and which flattens and smoothes out the Universe thus dispensing with the flatness and horizon problems.

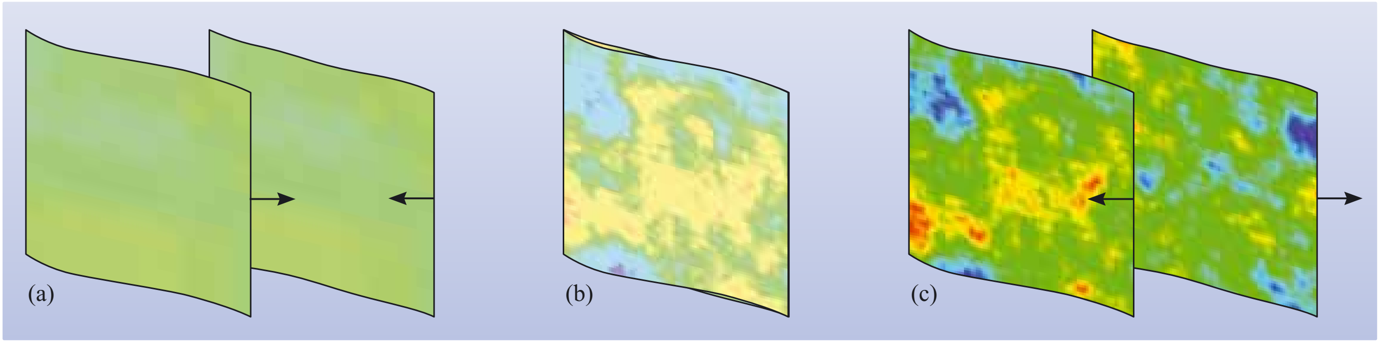


Figure 8.13 The ekpyrotic model, which invokes colliding sheets or 'branes' (representing our Universe and a parallel universe). (a) Two branes come together, resulting in (b) a collision that creates (c) the primordial fluctuations that are required to give rise to the observed large-scale structure and to the observed anisotropies in the cosmic microwave background.

In this theory the dark energy is a manifestation of the energy that controls the separation of the colliding branes, so it plays a vital role in the evolution of the cosmos rather than being a sort of 'uninvited guest'.

No one seems to know what the 'M' of M-theory stands for, though one of the theory's leading protagonists, Edward Witten, suggests 'magic, mystery or membrane, according to taste'.

8.8 The anthropic Universe

There seem to be a multitude of possible universes, but only one Universe. Why is our Universe the way it is? (This was the last question posed at the beginning – Problem 8.) Why do the fundamental constants of nature have the values they do? Could they have been different? In particular, if the Universe had started out with different conditions would we, or someone like us, still be here to puzzle about it?

The notion that the Universe must be able to give rise to life (and cosmologists!), has been given the status of a principle, the **anthropic principle**. It comes in two basic forms, the weak and the strong, and there are various versions of each.

The *weak anthropic principle* holds that the initial conditions of the big bang were such as to allow the eventual emergence of carbon-based life. Universes with different initial conditions – perhaps different values of the speed of light, the Planck constant, the gravitational constant, and so on – may have been possible, but if those conditions were not favourable to life then we would not be here. We would not, therefore, expect to find anything in the Universe incompatible with our own existence. So far the anthropic principle is not saying anything remarkable.

But the *strong anthropic principle* goes further, and holds that the Universe *necessarily* had the initial conditions that eventually allowed carbon-based life to emerge. Proponents of the strong principle argue that of all possible universes our Universe is so extraordinarily improbable that it cannot be an accident. Some go further and hold that the Universe was in some sense destined to develop not only life but self-awareness. They point to numerous coincidences that seem to imply that the Universe is finely tuned to favour the emergence of life.

The classic example, which actually pre-dates the anthropic principle, is the nucleosynthesis of carbon. Carbon is produced in post main sequence stars by the triple-alpha reaction. Two helium nuclei fuse to form an unstable nucleus of beryllium. Before the beryllium can fall apart again it is hit by a third helium nucleus to form a stable carbon nucleus. The trouble is, such a triple collision is so improbable that it's hard to explain the amount of carbon seen in stars today. The puzzle was solved in 1953 by the cosmologist Fred Hoyle, who predicted that carbon must possess a 'resonant' state such that at a certain collision energy a carbon nucleus is formed much more readily than expected. Without Hoyle's resonant state – which was later confirmed by experiment – there would be no carbon and no cosmologists. The energy of the resonant state is so finely tuned that if it were a little higher, the carbon would be rapidly converted into oxygen and there would again be no cosmologists. Hoyle himself was deeply affected by his discovery.

There are many other supposed coincidences of this nature though, unlike Hoyle's prediction, they have all been recognized in hindsight. Stable orbits are only possible in a universe with three spatial dimensions. Any more than three and there would be no planets for life to make its home on and no atoms either. The initial expansion of the big bang must have been just right – any faster and galaxies and stars would not have been able to form, any slower and the Universe would have collapsed again before life could emerge. Gravity must also be the right strength. If G is too big, only massive stars will form and burn out before life can take hold. If G is too small, stars will not get hot enough to start nuclear reactions.

Many people are unimpressed by such arguments, pointing out that no one should be surprised that the Universe permits us to exist. For them the anthropic principle begins and ends with its weak form and is not saying anything very profound. Douglas Adams, the author of the *Hitchhiker's Guide to the Galaxy*, liked to tell the story of a puddle of water lying in the road. Suppose the puddle suddenly becomes conscious and starts to contemplate its situation. It starts to sense its surroundings, probing the surface of the road beneath it, and notes that the depression in which it lies is the same shape as its own body. It comes to the conclusion not only that the Universe is perfectly adjusted to the emergence of puddles but that its own existence was somehow predestined. Can it be a coincidence?

You win the National Lottery against odds of 14 million to one. What an extraordinarily improbable coincidence! Yet someone had to win, and your win has to be seen in the context of millions of ticket holders who did not win. On this view, our Universe may one of countless possible past and future universes – perhaps arising from chaotic inflation – the difference being that ours holds the winning ticket.

Others argue that aside from the meaninglessness of hypothetical and unobservable 'other universes', such critics are missing the point. Russell Stannard, a physicist at the Open University, cites a counter-example of a prisoner sentenced to be executed by firing squad. At the crucial moment all ten marksmen miss their target and the prisoner is reprieved. Asked to explain his amazing deliverance the prisoner is unimpressed. 'Of course they missed, or else I wouldn't be here.' Such an explanation is not wholly satisfactory, since it fails to address *why* all ten skilled marksmen so improbably missed their target. Likewise, some people seek an explanation for why the Universe is set up the way it is beyond the explanation that it just has to be that way or else we would not be here to ask the question.

QUESTION 8.4

What do you think about the anthropic principle? Do you find it trivial, like Douglas Adams, or profound, like Russell Stannard? Can you find flaws in either of their stories? Can you reconcile the two views? How much is your opinion coloured by your own philosophical or religious beliefs? These are questions you will not find answered at the back of the book!

8.9 Epilogue

This chapter started with a quotation so it will finish with one too. Albert Einstein, whose insights into the nature of space and time made modern cosmology possible, once remarked that 'the most incomprehensible thing about the Universe is that it is comprehensible.' Perhaps after reading about multidimensional sheets banging together in 11-dimensional space, you may be inclined to think that Einstein, on this occasion, got it wrong!

But stay with us. Einstein was saying two things. First, he was expressing a faith that underpins all science, not just cosmology, namely that we will be able to understand the Universe and find it makes sense. The world is not chaotic and that makes science possible. But he was also saying something else, namely that we really have no right to expect the Universe to be that way. Why *should* we find the Universe comprehensible?

This could be a cue for another excursion into the anthropic principle, but we shall not do that. The models we have been discussing in the last few chapters may stretch your imagination to the limit (and beyond!) but they are the cosmologist's way of making the Universe comprehensible. So far our models have been able to keep up with new surprises sprung on us by the Universe. One day our luck and our imagination may run out and we may then have to admit that the Universe makes no sense after all. Until that happens, cosmology will continue to be one of the most exciting and mind-stretching of all the sciences.

QUESTION 8.5

Make brief notes *in your own words* to answer each of the questions posed at the beginning of this chapter. How satisfied are you with the answers? Check them against the summary at the end of this chapter.

8.10 Summary of Chapter 8

The models devised by cosmologists are simplified representations of the Universe. Like all models, they are only partial analogies to reality and break down outside their limits of validity. The big bang model is successful as far as it goes, but there are several problems it cannot answer.

- *Problem 1: What is the dark matter?* Dark matter makes up about 23% of the Universe. A little of it is baryonic, in the form of MACHOs, which are simply familiar objects that are too faint to see. Some of them can be revealed by

gravitational microlensing. About 85% of the dark matter has to be non-baryonic but apart from a very small proportion of neutrinos its nature is largely unknown. The best candidate is the neutralino, a form of WIMP, which may soon be discovered in laboratory experiments.

- *Problem 2: What is the dark energy?* Dark energy is a source of negative pressure that fills the Universe and drives the accelerating expansion. It should not be confused with dark matter. Its nature is still a mystery, but the leading contenders are Einstein's cosmological constant (a source of 'repulsive' gravity arising from general relativity), quantum vacuum energy (a consequence of Heisenberg's uncertainty principle) or 'quintessence' (an exotic form of matter).
- *Problem 3: Why is the Universe so uniform?* This is the horizon problem, which asks why widely separated regions have the same temperature and density, even though each has been beyond the horizon of the other throughout the history of the Universe. Inflation provides a possible answer. A small region of the Universe that had become homogeneous might have expanded so rapidly and by such an enormous factor that the whole of the currently observable part of the Universe (and perhaps more) is contained within the inflated homogeneous region.
- *Problem 4: Why does the Universe have a flat ($k = 0$) geometry?* Again, inflation may make it so. During the inflationary period large amounts of matter and energy were released into the Universe from the vacuum energy, leaving its density very close to the critical density, which corresponds to a flat geometry. Equivalently, whatever curvature the early Universe may have had would have been smoothed out by inflation leaving the spatial geometry of the observable Universe indistinguishable from that of a 'flat' space.
- *Problem 5: Where did the structure come from?* Clusters of galaxies were formed from density fluctuations in the early Universe which have left their imprint on the cosmic background radiation. Those fluctuations in turn may have arisen from tiny quantum fluctuations which were stretched by inflation from the microscopic scale up to and beyond the size of the then-observable Universe. At that point they would have become 'frozen in' as large-scale primordial fluctuations from which galaxies could condense.
- *Problem 6: Why is there more matter than antimatter?* Although one might expect equal numbers of particles and antiparticles to have been created in the early Universe, grand unified theories of physics allow a slight imbalance of matter over antimatter of 1 part in 10^9. The matter now in the Universe is that left over when the bulk of the matter and antimatter annihilated.
- *Problem 7: What happened at $t = 0$?* It's still anyone's guess. General relativity breaks down at the Planck time of 10^{-43} s, and to progress to earlier times requires a theory of quantum gravity that unifies general relativity with quantum physics. Inflation, too, remains without a firm grounding until this very early era is better understood. Limited progress has been made with quantum cosmology but the new all-encompassing M-theory offers several intriguing lines of enquiry.
- *Problem 8: Why is the Universe the way it is?* According to the anthropic principle, because we are here to ask the question!